

THE ELUSIVE TEN-SQUARE

G. HARDY ROPES

Goldens Bridge, New York

Hats off to Jeff Grant! In "9x9 Word Squares" in the November 1980 issue of Word Ways, he has rescued the nine-square from near-oblivion and done so within the confines of dictionaries. His article rekindled the spark of an old interest of mine -- the ten-square. For some time, I have felt (as Jeff Grant does) that constructing a ten-square is not beyond the realm of possibility.

During a brief visit with Palmer Peterson about four years ago, we talked about the difficulty of making various sizes of squares. He said he'd found an eight-square to be about 5 times as hard as a seven-square, and a nine-square again 5 times as hard as an eight-square. He supposed, therefore, that a ten-square would be another 5 times as tough as a nine-square. I don't think Palmer ever made an all-out effort to conquer the ten-square -- at least there is no evidence to that effect. (We were talking, of course, about non-tautonymic squares, not discounting the beauty of the tautonymic ones made in the past from words such as *ilang-ilang*, *Pangopango*, and the like.)

Murray Pearce of Bismarck, North Dakota later told me that Palmer Peterson's reason for being pessimistic about constructing a non-tautonymic ten-square was the fact that no comprehensive listing of ten-letter words has ever been made. By contrast, during the first third of this century an enormous amount of effort went into compiling nine-letter word lists in which words were indexed in reverse alphabetical order, so arranged because it is more efficient to build a large square from the bottom up rather than the top down. About 500,000 words were eventually listed, many drawn from reference works of the late nineteenth century.

One of the few sources of reverse-listed ten-letter words was compiled by the University of Pennsylvania in the 1960s under a contract with the Air Force Office of Scientific Research. They produced the Normal and Reverse English Word List, a listing taken mainly from Webster's Second, with 268,000 words of all lengths. Of this total, 35,200 (13 per cent) of the words were ten letters in length. (The only larger per centage was 13.6, for nine-letter words.) Unfortunately, plurals and other derived forms (-ed, -ing, -est) were not listed unless they were explicitly given in boldface in Webster. If they had been, the number of ten-letter words might have reached 50,000, still only one-tenth of the nine-letter corpus.

Given these facts, I propose the following strategies to make progress on the ten-square:

STRATEGY 1 Select a base word (last word in the square) that enables one to make maximum use of the far larger nine-lists, eight-lists, and seven-lists; such a word might contain Ss (plural words), Gs (-ing words) or Ds (-ed words).

STRATEGY 2 Locate a set of words for the seventh, eighth, ninth and tenth positions that provide relatively common endings for the first six words. What are the most common endings among ten-letter words? Data from nine-lists can help provide an answer. My Supplementary 9-List (compiled in 1932-34 by Rufus T. Strohm, Palmer Peterson and others) contains about 125,000 words; the principal source was the Oxford English Dictionary, but listings from four atlases and eight other references are also included. The most common four-letter endings in this list are -iver (2200), -ings (1500), -ling (700), -ngen (400), -esse (400), -berg (400), -iest (400), -ness (300), -ting (300), -ries (300), -lles (300), -ines (300), -ites (300) and -ters (300). Although analysis of other nine-lists might give slightly different results, it is highly probable that the most preferable terminal letters would remain S, R, G and E.

If these two strategies were sufficient, one or more ten-squares would have long ago been created by Palmer Peterson or the other great formists of the past. What else is needed?

STRATEGY 3 Codify rules concerning the admissibility of ten-letter words. Such rules were never, to my knowledge, clearly or consistently established by the early formists. It is apparent that for the ten-square criteria somewhat less stringent than Jeff Grant's must be allowed -- in particular, words or two-word phrases found in many different readily-available reference works (dictionaries, atlases and gazetteers, name directories). For example, I propose

- 1) Form words from the names of roads, lanes, streets or avenues listed in standard road maps: VALLEY ROAD, HARPER LANE, PARK STREET
- 2) Form words from place names in the United States followed by their state: ORONO MAINE, CANTON OHIO
- 3) Allow words from surnames in telephone directories of the major United States cities (readily available in libraries): BALLANTINE, STRICKLING
- 4) Allow words from first names and surnames from these directories: THOMAS HILL, AARON SMITH, HARDY ROPES

This would provide thousands of additional words; even more important, we could add many new common four-letter endings to the list, such as -road, -lane, -reet, -mith (for Smith), etc.

STRATEGY 4 Conquering the ten-square needs a concerted effort, not isolated work by a single individual; I call for a team effort. Some of

the steps that could be taken are: pooling of existing nine-lists, circulation of copies of existing ten-lists, location of other ten-lists; exchange of information about sets of basewords with good potential; other breakthroughs for new sources of words.

STRATEGY 5 If one or more readers of Word Ways have access to a digital computer even faster progress could result. However, two further preliminaries are necessary:

- 1) Conversion of the lists from Strategy 4 into a machine-readable form, a tedious and lengthy task. This work has already been done for the aforementioned Air Force List, as well as for a list of approximately 60,000 ten-letter words from Webster's Second and Third (plus derived forms) compiled by Jack Levine for his pattern word dictionaries. Does anyone have access to this latter computer tape?
- 2) Construction of an efficient algorithm for finding sets of words satisfying Strategy 2. Even with only 35,000 words, it is too much to ask the computer to look at all possible sets of four words; ingenious programming is needed to eliminate unpromising sets rapidly.

As Doug McIlroy demonstrated in the May 1976 Word Ways for the double six-square, a clever algorithm can rapidly discover a large number of word squares previously unknown to logologists. I have not seen the algorithm used by Frank Rubin to construct his near-miss ten-square in the February 1977 Word Ways, but it evidently did not use the bottom-up technique preferred by experienced formists.

I have the feeling that the conquest of the ten-square will be like the four-minute mile -- a seemingly impassable barrier, but, once down, bested on numerous occasions. I'd appreciate ideas and reactions sent to PO Box 533, Goldens Bridge NY 10526. Climb on the bandwagon!